

**Univerzita Karlova v Praze**

**1. lékařská fakulta**

Autoreferát disertační práce



Extrakce informací z lékařských textů  
Extracting Information from Medical Texts

Ing. Karel Zvára

2017

## **Doktorské studijní programy v biomedicině**

*Univerzita Karlova v Praze a Akademie věd České republiky*

Obor: Biomedicínská informatika

Předseda oborové rady: prof. MUDr. Štěpán Svačina, MBA, DrSc.

Školící pracoviště: Vysoká škola ekonomická v Praze

Školitel: doc. Ing. Vojtěch Svátek, Dr.

Konzultant (byl-li):

Disertační práce bude nejméně pět pracovních dnů před konáním obhajoby zveřejněna k nahlížení veřejnosti v tištěné podobě na Oddělení pro vědeckou činnost a zahraniční styky Děkanátu 1. lékařské fakulty.

## Obsah

Abstrakt.....	4
Abstract.....	5
Klíčová slova .....	6
Keywords .....	6
1. Úvod.....	7
2. Hypotézy a cíle práce.....	7
3. Materiál a metodika .....	8
4. Výsledky .....	12
5. Diskuse.....	15
6. Závěry .....	17
7. Použitá literatura .....	18

## Abstrakt

Cílem mé práce bylo zjistit specifické vlastnosti českých lékařských zpráv z hlediska možnosti extrahovat z nich konkrétní informace.

Pro svoji práci jsem měl k dispozici celkem 268 anonymizovaných narativních lékařských zpráv ze dvou ambulantních pracovišť. Studoval jsem standardy pro uchování elektronické zdravotnické dokumentace i pro přenos klinických informací mezi informačními systémy ve zdravotnictví. Věnoval jsem se také implementování elektronického zdravotního záznamu v zubním lékařství.

Nejprve jsem se narativní lékařské zprávy snažil zpracovat pomocí nástrojů pro zpracování přirozeného jazyka (Natural Language Processing, NLP). Dospěl jsem k závěru, že narativní lékařské zprávy v českém jazyce jsou typickému českému textu velmi vzdálené zejména pro svoji heslovitost a absenci české větné stavby. Obsahují také velké množství překlepů, zkratk a zkrácených slov. Vzhledem k nedostupnosti hlavních mezinárodních klasifikačních systémů v českém jazyce jsem se rozhodl pokračovat ve výzkumu vývojem metody pro přípravu vstupního textu pro překlad a jeho sémantickou anotaci.

Hlavním cílem této části výzkumu bylo navrhnout metodu a podpůrný software pro interaktivní korekci a sémantickou anotaci narativních lékařských zpráv, které by umožnily jejich snadnější použití, s menším množstvím chyb i mimo jejich původní kontext.

Vyvinul jsem třífázovou metodu předběžného zpracování s cílem podpořit druhotné využití lékařských zpráv. Metoda třífázového předzpracování narativních klinických zpráv byla ověřena na 49 anonymních českých lékařských zprávách z oblasti kardiologie.

## Abstract

The aim of my work was to find out the specific features of Czech medical reports in terms of the possibility of extracting specific information from them.

For my work, I had a total of 268 anonymized narrative medical reports from two outpatient departments. I have studied standards for preserving electronic health records and for transferring clinical information between healthcare information systems. I have also participated in the process of implementing electronic medical record in the field of dentistry.

First of all, I tried to process narrative medical reports using natural language processing (NLP) tools. I came to the conclusion that narrative medical reports in the Czech language are very different than a typical Czech text, especially because it mostly contains short telegraphic phrases and the texts lack typical Czech sentence structure. It also contains many misspellings, acronyms and abbreviations. Another problem was the absence of existence of the Czech translation of the main international classification systems. Therefore I decided to continue the research by developing the method for pro-processing the input text for translation and its semantic annotation.

The main objective of this part of the research was to propose a method and support software for interactive correction and semantic annotation of narrative medical reports that would allow their easier use with fewer errors outside of their original context.

I have developed a three-phase pre-processing method to encourage secondary use of medical reports. The method of three-phase pre-processing of narrative clinical reports was verified on 49 anonymous Czech medical reports in the field of cardiology.

## **Klíčová slova**

Lékařské zprávy, zpracování přirozeného textu, extrakce informací, elektronický zdravotní záznam.

## **Keywords**

Clinical reports, natural language processing, information extraction, electronic health record.

# 1. Úvod

Lékařské texty, kterými se v práci zabývám, jsou texty, které vznikají v průběhu poskytované zdravotní péče v diagnosticko-terapeutickém procesu. Patří k nim zejména lékařské zprávy, které jsou významnou součástí zdravotnické dokumentace. Slouží pro uchování zjištěných informací o zdraví, o provedených úkonech, léčbě a souvisejících administrativních údajích. V České republice mají lékařské zprávy obvykle povahu volného textu, který je formátován jen pomocí mezer, tabulátorů a nových řádků. Lékařské zprávy, jako součást zdravotnické dokumentace obsahují zejména identifikační a administrativní údaje, údaje o zdravotním stavu či okolnostech úmrtí, rozhodnutí a pokyny ošetřujícího lékaře, případně další zákonem vyžadované údaje.

Zdravotnická dokumentace neslouží jen pro uchování informací pro další poskytování zdravotní péče. Je také důležitým důkazem o vyhovění státním požadavkům, o řádné péči ve vztahu k pacientovi a také pro vyúčtování nákladů na péči.

## 2. Hypotézy a cíle práce

Jako cíl práce jsem na počátku studia stanovil „*zjištění specifických vlastností českých lékařských zpráv z hlediska možnosti extrahovat z nich konkrétní informace*“. Očekával jsem, že pro realizace tohoto cíle je nutné zjistit, které vlastnosti českých lékařských zpráv působí největší problémy v nestatických fázích zpracování přirozeného jazyka, navrhnout základní postup pro analýzu česky psaných lékařských zpráv a pomocí vlastní implementace postup ověřit a výsledky publikovat.

Nejprve jsem se věnoval snahám o analýzu dostupných narativních českých lékařských zpráv pomocí nástrojů NLP. Navázal jsem

především na práci Jiřího Semeckého, Petera Smatany a Petry Přečkové.

Inspiroval jsem se také přístupy použitými v zahraničí na anglicky psaných narativních lékařských zprávách, zejména možností jejich předzpracování.

### 3. Materiál a metodika

#### **Materiál**

Pro svoji práci jsem měl k dispozici celkem 268 anonymizovaných narativních lékařských zpráv v českém jazyce. Tyto zprávy byly zbaveny identifikačních údajů a pocházely ze dvou ambulantních kardiologických pracovišť Městské nemocnice Čáslav z období let 2000 až 2004. Tyto zprávy byly získány s informovaným souhlasem pacientů.

#### **Klasifikační systémy a další standardy**

Zdravotnická dokumentace je v elektronické podobě ukládána a přenášena mimo jiné s využitím národních i mezinárodních standardů. Mezinárodní standardy pro ukládání elektronické zdravotnické dokumentace pocházejí zejména od standardizačních organizací Health Level Seven Inc. a IEEE. Standardy ve spolupráci zejména s Health Level Seven Inc. publikuje také ISO a CEN.

Rozlišují se standardy pro uchování elektronické zdravotní dokumentace (clinical document) a standardy pro přenášení údajů ve zdravotnictví (messaging). Přenášení klinických údajů se věnují zejména standardy HL7 verze 2 a HL7 verze 3, jejichž scénáře použití dále stanovuje sdružení Integrating Healthcare Enterprise (IHE). Pro uchování zdravotnické dokumentace se zejména v USA používá standard Continuity of Care Document (CCD) pocházející od Health Level Seven Inc. Obecnou standardizaci poskytl standard ISO EN 13606, (EHRcom) který se věnuje jak uložení klinických a



administrativních informací, tak i jejich předávání. Reflektuje přitom existenci dalších standardů Health Level Seven i openEHR používaného zejména v Austrálii. V Evropě byl v rámci projektu epSOS vytvořen standard epSOS Patient Summary, který definuje strukturu minimálního elektronického zdravotního záznamu nutného pro přeshraniční péči.

Strukturovaná elektronická zdravotnická dokumentace i její předávání vyžadují využití klasifikačních systémů pro kódování hodnot. Mezinárodní standardy využívají zejména standardy LOINC (Logical Observation Identifiers Names and Codes), SNOMED CT (Systematized Nomenclature of Medicine Clinical Terms) a Mezinárodní klasifikaci nemocí. Tyto klasifikační systémy, zejména SNOMED CT, jsou založené na vzájemně provázaných konceptech. Využití těchto klasifikačních systémů předpokládají i standardy EHRcom či epSOS Patient Summary.

V České republice se z mezinárodních standardů používá zejména HL7 v2, a to pro komunikaci klinických systémů s přístroji. Pro předávání klinických údajů mezi nemocnicemi (i hlášení do zdravotnických registrů) se využívá český standard Dasta (Datový standard MZ ČR). Pro kódování hodnot se obvykle využívá soustava číselníků *Národní číselník laboratorních položek* (NČLP).

Jednotlivé informační systémy používají zpravidla interní struktury pro zachycení klinických a administrativních údajů. V některých případech se kromě narativních zpráv využívá také strukturovaná zdravotní dokumentace, a to zejména v zubním lékařství.

### **Zpracování přirozeného textu**

Volný text je nejprve zapotřebí připravit pro další zpracování. Základní metodou je tokenizace vstupního textu, tedy jeho rozdělení do řetězce strojově oddělených součástí. Každá taková součást obvykle sestává z posloupnosti znaků, jednotlivé součásti (tokeny)

jsou odděleny tzv. stop-znaky, kterými bývají mezery, pevné mezery, tabulátory či konce řádků.

Na tokenizaci obvykle navazuje stemizace či lemmatizace, což jsou metody, kterými lze hledat kořen slova a jeho další vlastnosti ukryté v jeho konkrétní formě. Stemizace je založena na nalézání kořene slova pomocí odstraňování předpon a přípon. Lemmatizace je opačným přístupem, kdy se ze slovníku pomocí pravidel pro tvorbu a ohýbání slov vytvářejí jednotlivé tvary, které se následně vyhledávají v textu.

### **Předzpracování textu (pre-processing)**

Pro předzpracování je využito tokenizace, podobně jako v případě zpracování zpráv nástroji pro zpracování přirozeného jazyka.

Tokenizovaná zpráva následně prochází fázemi normalizace a sémantické anotace. Obě fáze vyžadují aktivní účast odborného lékaře, který je schopný porozumět vstupním narativním lékařským zprávám.

Fáze normalizace slouží pro opravu původní narativní textové zprávy pomocí pěti nástrojů:

1. rozdělení nesprávně spojených slov
2. spojením nesprávně rozdělených slov
3. mazání nesprávně vložených slov
4. změna obsahu slov z důvodu překlepů
5. změna obsahu slov z důvodu rozvinutí zkratky nebo zkráceného slova

Fáze sémantické anotace slouží k označení příslušnosti konkrétních termínů normalizované zprávy k termínům klasifikačních systémů a určení vztahu zmínky k přítomnosti příslušného stavu. V lékařských zprávách jsou totiž obsažené informace nejen o přítomnosti konkrétního stavu, ale mohou v nich být obsažené také informace o jejich absenci.

Výsledkem třetí fáze je částečně strukturovaná normalizovaná lékařská zpráva sestávající ze dvou částí. První částí je normalizovaná klinická zpráva s využitím strojového překladu, která je snadno čitelná i pro lékaře neznalé českého jazyka. Druhou část tvoří souhrn strukturované informace nalezené v normalizované lékařské zprávě.

Pro třífázovou metodu předzpracování lékařských zpráv (tokenizace, normalizace, sémantická anotace) jsem vyvinul softwarový nástroj TOCESA.

Použitím nástroje vzniká normalizační databáze, z ní lze odhadnout potenciál pro učení automatizované transformační procedury, tedy zda dostatečný počet nezpracovaných termínů vždy transformuje buď na stejný termín, nebo na skupinu termínů, které mají jen nepatrné gramatické rozdíly.

Odhad míry chyb lze vyjádřit takto:

$$Err = \frac{\sum_r (n_r - 1) \cdot w_r}{\sum_r (n_r - 1)}$$

Kde  $n_r$  je počet výskytů nezpracovaného výrazu  $r$  a  $w_r$  je váha  $r$  pro výpočet chyby a váha se zjistí jako:

$$w_r = \frac{|\{(i, j)\}; i, j \in tsf(r), i \neq j|}{P(|tsf(r)|, 2)}$$

Kde  $tsf(r)$  je množina transformovaných pojmů, které jsou výsledkem nezpracovaného výrazu  $r$  v databázi,  $i$  a  $j$  jsou výrazy z této množiny (jsou jeho různé prvky, ačkoli možná stejné termíny) a  $P(n, k)$  označuje  $k$ -permutaci  $n$ .

## 4. Výsledky

### **Zpracování nástrojů pro zpracování přirozeného jazyka**

Rozpoznávací mechanismus pro vyhledání kombinací odpovídajících krevnímu tlaku našel celkem 434 případů, přičemž minimum počtu nalezených kombinací ve zprávě bylo 0, nejvyšší počet rozpoznaných zápisů krevního tlaku bylo 12.

Pomocí slovníku iSpell a využitím nových pravidel pro odvozování tvarů slov pomocí ohýbání (lemmatizace) se podařilo identifikovat jen 47,7 % tokenů.

Číselníkové termíny se podařilo nalézt v celkem 107 zprávách. Těchto 107 zpráv bylo automatizovaně rozděleno na celkem 66 376 tokenů. Až na výjimky se podařilo rozpoznat jen termíny z české verze bibliografického klasifikačního systému MeSH. Termíny ze SNOMED CT se nedařilo rozpoznat, neboť nebyla k dispozici česká verze SNOMED CT. Snaha o rozpoznání položek z číselníku mezinárodní klasifikace nemocí MKN10 nevedla k velkému počtu rozpoznání.

### **Ukládání strukturovaného zdravotního záznamu**

V rámci výzkumu jsem navrhl metodu a vytvořil softwarové řešení zahrnující práci s interaktivní komponentou zubního kříže Lifetime DentCross pro celoživotní elektronický zdravotní záznam ve stomatologii, který zahrnuje možnost vkládání údajů nejen pro stálý, ale i pro smíšený a dočasný chrup. Pro tyto situace byl rozšířen i model ontologie stomatologie 2. lékařské fakulty Univerzity Karlovy v Praze a Fakultní nemocnice v Motole a je popsán v disertační práci K. Chleboráda.

### **Předzpracování narativních lékařských zpráv**

Metoda 3PP byla aplikována na údaje 49 anonymních českých lékařských zpráv z oblasti kardiologie. Dva kardiologové metodu 3PP v software TOCESA ověřili v letech 2012 až 2015 na celkem 49 lékařských zprávách z oblasti kardiologie.

V první fázi bylo 49 anonymizovaných narativních lékařských zpráv rozděleno do celkem 3 324 tokenů (včetně mezer). V těchto 49 tokenizovaných zprávách provedla kardioložka celkem 148 oprav, při kterých ze 127 různých znění tokenů s překlepy vytvořila celkem 129 různých slov.

Ze 450 znění zkrácených slov bylo vytvořeno 684 slov či slovních spojení v celkem 1 411 případech. Z 92 různých znění zkratk bylo vytvořeno celkem 136 různých plných znění v celkem 267 případech. Zkrácená slova byla rozvinuta na různý počet slov po rozšíření, nejvíce na osm, zkratky nejvýše na pět slov.

Provedl jsem analýzu normalizační databáze s cílem zjistit její opětovnou použitelnost. V případě oprav chyb nebyly zjištěny žádné heterogenní transformace, míra chybovosti byla tedy 0 %. Podobně bylo zjištěno jen málo opakovaných korekcí stejné chyby (19 %), takže poměr transformace opětovaného použití je nízký.

Abych zjistil dopad normalizace na strojový překlad, aplikoval jsem překladač Google na původní i normalizovaný text. Vzhledem k tomu, že před odevzdáním disertační práce Google spustil novou verzi překladače, provedl jsem experiment také opakovaně v nové verzi překladače. Na obou níže uvedených obrázcích jsou rozdíly mezi strojovým překladem původní a normalizované zprávy.

### *Příklad pomoci původního překladače Google:*

Patients with implanted pacemaker Biotronik Pikos mode VVIM for AV block second-degree-burns-II. degree. Subj: he manages quite well against min.vyš not change. Vol-- no changes compared to the previous examination. Vol. : cooperates, Cooperates, weight 64 kg, height-height: 168 cm-cm, BMI: 25.25; 3. Head, neck: BPN--no positive findings, BP: 125/75, Thorax: symmetrical, percussion, clear, vesicular breathing, clean hearts: Tapping the mdel-ē.mdel. no. , heartbeat-Heartbeat regular at the apex and on the basi two darker sounds, belly: abdomen: soft, palp.nobel.palp. was not. , hepar, Hepar, lien: not enlarged, tapottement Bilat. neg. DK: no swelling, the calf was not. ECG: stimuleovaný-regular rhythm of-stimulated ventricular VVI mode - 100% stimulation. Parameters: Frequency base: 74-71, 7 / min.min. , Magnetic frequencies: 74-frequency: 71, 7 / min. interval: 836 ms pulse width of 0.width: 0, 72 ms Pacemaker Function: chol correct. 5, 3 mmol / l HDL 4-1, 1 mmol / l LDL 3-2mmol-3, 2 mmol / L TG 4-1, 5 mmol / l glykemie 6.6, 1 Conelusion-mmol--mmol / l Conclusion. : Implantaqe-Implantation of a permanent pacemaker VVI, M-mode VVIM for AV block second-degree-burns-II. st. sinus and bradycardia with the final events-resulting in action of 40 / min.min. , and-And PAUSE 3-5ms-3, 5 ms. HLP intervened statins CHD comp-compensated. Arterial hypertension III.st. degree. - well-Well controlled. Therapie-Therapie. : Anopyrin 400mg-100 mg: 1 / d, HCHTH 4-2tb-1/2 tb / d, Enap 5mg-1-0-1-5 mg: 1 - 0-1 / d, Zocor 20mg: 0-0-1-20 mg: 0-0 - 1 / d, control at-TK, blood pressure, minerals. The principles of secondary prevention, anti-sclerotic diet. Checking in 4 months

### *Příklad pomoci nové verze překladače Google:*

Patient with implanted pacemaker Biotronik Pikos in VVIM mode for AV block II. Degree. Subj: he's doing pretty well - there's no change against min. Obj- Collaborates, to the previous exam. Order no. : Cooperates, weight: 64 kg, height: 168 cm, BMI: 25.3-25, 3. Head, neck: bpn--no positive finding, TK: 125/75, chest: symmetrical, full throat, tap full, clear, breathing cellar, clean, heart: elue-tick to mdel.mdel. No. , Regular heart rate, bite Heart rate regular, on the tip and bass on the base 2 darker, Soft, palp.nobel darker echoes, abdomen: soft, palp, palp. It was not. , Hepar, lien: not enlarged, tapottement bilat. Neg. DK: no swelling, no calf. ECG: Regular stimulated-stimulated ventricular rhythm, VVI mode - 100% stimulation. Device parameters: Basic-Base frequency: 74-71, 7 / min.min. , Magnetic frequency: 74-71, 7 / min. Interval: 836 ms pulse width: 0-0, 72 ms Pacemaker functions: correct chol. 5, 3mmol-3 mmol / l HDL 4-1, 1 mmol / l LDL 3-3, 2 mmol / l TG 4-1, 5 mmol / l glycemia 6.6, 1 mmol Conclusion: Implantation of a permanent- / l Conclusion. : Continuous pacemaker implantation in VVI, M-VVIM mode for AV blockade-block II. St. And sinus bradycardia with a resulting action of 40 / min, and min. , And pauses of 3-5ms-3, 5 ms. HLP intervened-interfered with ICHS statins ICHS comp-compensated. Arterial hypertension III.st. Degree. - well controlled. Therapie- Therapie. : Anopyrin 400mg-100 mg: 1 / d, HCHTH 4-2tb-1/2 tb / d, Enap 5mg-1-0-1-5 mg: 1 - 0 - 1 / d, Zocor 20mg-0-0-1-20 mg: 0 - 0 - 1 / d-d. Principles of secondary prevention, antisklerotic diet. Check for 4 months

Je vidět, že dopad na překlad je stále značný, i přes podstatné vylepšení Google překladače.

Normalizované lékařské zprávy se lišily délkou, od té se odvíjel i počet sémantických anotací. Mezi 49 normalizovanými zprávami měla nejkratší jen 244 znaků bez mezer a nejdelší 3 153 znaků bez mezer. Průměrná zpráva zprávy byla 1 576 znaků bez mezer. První kardioložka provádějící sémantickou anotaci anotovala celkem 49 normalizovaných lékařských zpráv, přičemž zaznamenala celkem 1 500 anotací.

Druhý kardiolog následně provedl revizi sémantických anotací provedených první kardioložkou. V případě 1 454 anotací (96,96 %) se s ní shodl, dalších 46 anotací bylo anotováno jinak. (3,07 %)

## 5. Diskuse

### **Zpracování nástrojů pro zpracování přirozeného jazyka**

Narativní lékařské zprávy, které jsem měl pro účely výzkumu k dispozici, obsahovaly velké množství zkratk a zkrácených slov. Výsledky mého výzkumu tím jen potvrzují předchozí výsledky, kterých dosáhli Petra Přečková a další. Podařilo se rozpoznat jen malou část lékařských termínů a to téměř výhradně jen s využitím bibliografického, nikoliv klinického, klasifikačního systému. S pomocí techniky lemmatizace se navíc podařilo jako česká slova identifikovat méně než polovinu alfanumerických tokenů.

Ukázalo se, že narativní české lékařské zprávy nejsou typickým českým narativním textem. Jsou textem velmi specifickým, který je z velké části složen z odborných termínů a vyznačuje se intenzivním využitím zkratk, zkrácených slov a absencí explicitního vyjádření větných členů.

### **Ukládání strukturovaného zdravotního záznamu**

Kromě získání zkušenosti s konkrétním záznamem je zajímavým výsledkem poznatek o důležitosti samotné techniky zadávání údajů do elektronického zdravotního záznamu. Zubní lékař totiž během ošetřování nemůže ovládat výpočetní techniku pomocí rukou, neboť by tak kontaminoval rukavice, které používá při vyšetřování či ošetřování pacienta. Hlasový vstup lékaři umožňuje zaznamenávat údaje do elektronického zdravotního záznamu bez neustálé výměny rukavic či (dnes obvyklé) asistence dalšího pracovníka.

## **Předzpracování narativních lékařských zpráv**

Pokud lékař, který ošetřuje pacienta, plynule ovládá jazyk lékařské zprávy, problémy s jejím narativním textem nebývají kritické, i když i v tom případě se může potýkat s problémy s porozuměním zkratkám a zkráceným slovům jiných specializací. Strojový překlad takového původního textu, jehož cílem je přiblížit význam částí lékařské zprávy lékaři v cizím jazyce, pravděpodobně narazí na vážné potíže.

Společně se spolupracujícími kardiology jsme zjistili, že po dokončení první a druhé fáze metody 3PP bylo snadno čitelných všech 49 ze 49 zpracovaných lékařských zpráv. Ověření metody ukázalo, že nejnáročnější prací, kterou provedli spolupracující kardiologové, byla sémantická anotace prováděné ve 3. fázi metody 3PP. Důvody pro to byly tři:

1. Klasifikační systémy SNOMED CT a LOINC nejsou dostupné v českém jazyce.
2. Vyhledané klinické pojmy často neodpovídaly zapsanému údaji přesně, lékařům se podařilo najít jen buď konkrétnější, nebo obecnější termín.
3. Oba kardiologové se nejprve museli naučit software používat. Rychlost anotování se ziskem zkušeností vzrůstala. Třetí kardiolog, kterého jsme se pokusili do práce též zapojit, nebyl schopný naučit se software dostatečně ovládat.

Třífázová metoda předzpracování (3PP) má svá omezení:

1. Nástroj TOCESA není zatím dostatečně robustní pro široké využití a databáze transformací je zatím velmi malá a zkrácená skutečností, že zprávy pocházely od malého počtu lékařů.
2. Databáze anotací je závislá na stávajících klasifikačních systémech, které prozatím nejsou dostatečně propojené a konzistentní.



## 6. Závěry

Narativní lékařské zprávy umožňují lékařům sdílet složité představy o onemocnění pacienta a navrhované terapii a odrážejí pozorování lékaře, myšlení, hodnocení a důvody pro plán léčby. Jsou důležitou součástí zdravotnické dokumentace, nicméně jejich opětovné využití prostřednictvím informačních a komunikačních technologií je obtížné z důvodu mnoha chyb a nejednoznačností.

Snaha o lingvistickou analýzu na podkladových narativních lékařských zprávách potvrdila dřívější závěry Jiřího Semeckého a Petera Smatany ohledně povahy těchto zpráv, které nejsou typickými českými texty. Jde spíše o heslovitá sdělení s častým využitím zkratk a zkrácených slov a s velkým množstvím překlepů.

České lékařské zprávy nelze proto považovat za běžné české texty. Snaha o extrakci informací byla značně negativně ovlivněna absencí mezinárodních klasifikačních systémů jako je LOINC a SNOMED CT v českém jazyce.

V poslední části studia jsem se zaměřil na návrh třífázové metody pro předzpracování původních textů, aby bylo možné je lépe využít pro extrakci informací, například lékařem v zahraničí, který poskytuje neplánovanou péči. Tato navrhovaná metoda a její nástroje umožňují zvýšit kvalitu textu a provést sémantickou anotaci.

Provedené ověření, byť založené na malém vzorku, naznačuje, že provedení transformací termínů má kladný dopad na kvalitu strojového překladu. Obtížně pochopitelné (a v současné době strojově nepřeložitelné) termíny jsou nahrazené přiměřenými překlady. Odhadovaná kvalita plně automatizované transformace je slibná ve srovnání s odstraněním velké náročnosti ručního zpracování, i když transformace nejednoznačných tokenů bude zřejmě vyžadovat ruční zásah.

Otázku, kterou jsem si položil v pojednání o disertační práci: „*Které vlastnosti českých lékařských zpráv působí největší problémy v nestatistických fázích zpracování přirozeného jazyka?*“ jsem zodpověděl.

## 7. Použitá literatura

- [1] Semecký J, Zvárová J (vedoucí diplomové práce). Diplomová práce: Multimediální elektronický záznam o nemocném v kardiologii. Praha: Matematicko-fyzikální fakulta Univerzity Karlovy; 2001.
  
- [2] Smatana P, Paralič J (vedoucí diplomové práce). Diplomová práce: Spracovanie lekárskych správ pre účely analýzy a dolovania v textoch. Košice: Technická univerzita v Košiciach; 2005.
  
- [3] Přečková P, Zvárová J (školitel). Disertační práce: Jazyk lékařských zpráv a jeho informačně lexikální analýza. Praha: 1. lékařská fakulta Univerzity Karlovy; 2011.
  
- [4] Cheng TO. Medical Abbreviations.. J R Soc Med. 2004;97(11)
  
- [5] Van Ginneken AM. The physician´s flexible narrative. Methods Inf Med. 1996;35(2)
  
- [6] Van Ginneken AM. Considerations for the representation of meta-data for the support of structured data entry. Methods Inf Med. 2003;42

- [7] Van Ginneken AM., Stam H, Van Mulligen EM, de Wilde M, Van Mastriigt R, Van Bommel JH. ORCA: the versatile CPR. *Methods Inf Med.* 1999;38
- [8] Nagy M, Říha A (školitel). *Disertační práce: Harmonizace klinického obsahu elektronického zdravotního záznamu.* Praha: Univerzita Karlova v Praze; 2011.
- [9] Shapiro JS, Bakken S, Hyun S, Melton GB, Schlegel C, Johnson SB. Document Ontology: Supporting Narrative Document in Electronic Health Records. *AMIA Annu Symp Proc.* 2005;
- [10] Bleeker SE, Derksen-Lubsen G, Van Ginneken AM, Van der Lei J, Molla HA. Structured Data Entry for Narrative Data in a Broad Speciality: Patient History and Physical Examination in Pediatrics. *BMC Med Inform Decis Mak.* 2006;
- [11] Knaup P, Garde S, Haux R. Systematic planning of patient records for cooperative care and multicenter research. *Int J Med Inf.* 2007;76
- [12] Deleger L, Grouin C, Zweigenbaum P. Extracting medical information from narrative patient records - the case of medication/related information. *J Am Med Inform Assoc.* 2010;17
- [13] Garcia-Remesal M, Maojo V, Billhardt H, Crespo J. Integration of Relational and Textual Biomedical Sources. *Methods Inf Med.* 2009;48(1)

- [14] Blaschke C, Hirschman L, Valencia A. Information extraction in molecular biology. *Brifings in Bioinformatics*. 2010;2(1)
- [15] Johnson SB, Bakken S, Dine D, Hyun S, Mendonce E, Morrison F et al. An electronic health record based on structure narrative. *J Am Med Inform Assoc*. 2008;15(1)
- [16] Hui Y. Automatic extraction of medication information from medical discharge summaries. *J Am Med Inform Assoc*. 2010;17
- [17] Meystre S, Haug PJ. Natural language processing to extract medical problems from electronic clinical documents. Performance evaluation.. *J Biomed Inform*. 2006;39
- [18] Oemig F, Blobel B. Natural Language Processing Supporting Interoperability in Healthcare. *Text Mining*. 2014;
- [19] Lopprich K, Krauss F, Ganzinger M, Senghas K, Riezler S, Knaup P. Automated classification of selected data elements from free/text diagnostic reports in clinical research. *Methods Inf Med*. 2016;55
- [20] Meystre SM, Savova GK, Kipper-Schuler KC, Hurdle JF. Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research. *IMIA Yearbook of Medical Informatics*. 2008;
- [21] Friedman C, Shagina I, Lussier I, Hripscak G. Automated encoding of clinical documents based on natural language

- processing in the clinical environment. J Am Med Inform Assoc. 2004;
- [22] Friedman C. System and method for language extraction and encoding utilizing the parsing of text data in accordance with domain parameters. United States Patent no. 6,182,029 B1. USA. 2001
- [23] Riskin D, Shroff A. Systems and methods for processing patient data history. United States Patent no. 2014/0181128 A1. USA. 2014
- [24] Safran C, Bloomrosen M, Hammond W, Labkoff S, Markel-Fox S, Tang P et al. Toward a National framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper. J Am Med Inform Assoc. 2007;14(1)
- [25] Schreier G, Ammenwerth , Hörbst A, Hayn D. eHealth2016 - Health Informatics Meets eHealth. IOS Press Ebook. 2016
- [26] Dostál O, Šárek M. Support for Electronic Health Records in Czech Law. Eur J Biomed Inform. 2012;8(2)
- [27] Blobel B, Giacomini M. Interoperability is more than just technology. Eur J Biomed Inform. 2016;12(1)
- [28] Žďárek R. Vedení zdravotnické dokumentace a její náležitosti. Zdravotnické noviny. 2009

- [29] Zákon o péči o zdraví lidu. In: Sběrka zákonů. Ministerstvo spravedlnosti ČSSR; 1966. p. 74-91.
- [30] Novela zákona o péči o zdraví lidu. In: Sběrka zákonů. Tiskárna Ministerstva vnitra, p.o.; 2001. p. 6344-6346.
- [31] Vyhláška o zdravotnické dokumentaci. In: Sběrka zákonů. Tiskárna Ministerstva vnitra, p.o.; 2006. p. 5282-5284.
- [32] Zákon o zdravotních službách a podmínkách jejich poskytování. In: Sběrka zákonů. Tiskárna Ministerstva vnitra, p.o.; 2011. p. 4730-4801.
- [33] Hammond K., Helbig S., Benson C., Brathwaite-Sketoe B.. Are Electronic Medical Records Trustworthy? Observations on Copying, Pasting and Duplication.. AMIA Annual Symposium Proceedings. 2003;
- [34] HL7 Standards - Primary Standards, available from [http://www.hl7.org/implement/standards/product\\_section.cfm?section=1](http://www.hl7.org/implement/standards/product_section.cfm?section=1) (cited 19.5.2017)
- [35] The DICOM Standard, available from <http://dicom.nema.org/standard.html> (cited 19.5.2017)
- [36] Health Level Seven HL7 version 3 - Reference Information Model. International Organization for Standardization; 2014.
- [37] Who is IHE?, available from [http://www.ihe.net/FAQ/#Who\\_is\\_IHE?](http://www.ihe.net/FAQ/#Who_is_IHE?) (cited 19.5.2017)

- [38] Datový standard MZ ČR, available from  
<http://ciselniky.dasta.mzcr.cz/> (cited 19.5.2017)
- [39] Unified Medical Language System (UMLS), available from  
<https://www.nlm.nih.gov/research/umls/> (cited 20.5.2017)
- [40] About epSOS, available from  
<http://www.epsos.eu/home/about-epsos.html> (cited 26.4.2017)
- [41] ASTM E2369-12 Continuity of Care Record. American Society for Testing and Materials;
- [42] Häyrinenn K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *Int J Med Inf.* 2008;5
- [43] What is openEHR?, available from  
[http://www.openehr.org/what\\_is\\_openehr](http://www.openehr.org/what_is_openehr) (cited 19.5.2017)
- [44] CORDIS: The Good European Health Record, available from  
[http://cordis.europa.eu/project/rcn/17093\\_en.html](http://cordis.europa.eu/project/rcn/17093_en.html) (cited 19.5.2017)
- [45] Origins of openEHR, available from  
<http://www.openehr.org/about/origins> (cited 19.5.2017)
- [46] Archetype Technology Overview, available from  
<http://www.openehr.org/releases/AM/latest/docs/Overview/Overview.html> (cited 19.5.2017)

- [47] ISO EN 13606. International Organization for Standardization / European Committee for Standardization; 2008.
- [48] ISO EN 13940. International Organization for Standardization / European Committee for Standardization; 2015.
- [49] epSOS: Patient Summary, available from <http://www.epsos.eu/epsos-services/patient-summary.html> (cited 19.5.2017)
- [50] D3.2.2 Final epSOS System Technical Specification, available from [http://www.epsos.eu/uploads/tx\\_epsosfileshare/D3.2.2\\_Final\\_Definition\\_Functional\\_Service\\_Req\\_Patient\\_Summary.pdf](http://www.epsos.eu/uploads/tx_epsosfileshare/D3.2.2_Final_Definition_Functional_Service_Req_Patient_Summary.pdf) (cited 19.5.2017)
- [51] NIX-ZD, available from <http://www.nix-zd.cz/> (cited 20.5.2017)
- [52] Situace kolem registrů zavedených zákonem o zdravotních službách, available from <http://www.sasp.cz/situace-kolem-registru-zavedenych-zakonom-o-zdravotnich-sluzbach> (cited 20.5.2017)
- [53] Přečková P. Jazyk českých lékařských zpráv a klasifikační systémy v medicíně. Eur J Biomed Inform. 2010;6(1)
- [54] Get LOINC, available from <https://loinc.org/downloads/> (cited 03.06.2017)



- [55] International Classification of Diseases v. 10 (ICD10), available from <http://www.cdc.gov/nchs/> (cited 19.5.2017)
- [56] International language for drug utilization research, available from <https://www.whocc.no/> (cited 3.6.2017)
- [57] Converging Patient Summaries: Finding the Common Denominator between the European Patient Summary and the US-Based Continuity of Care Document. *Eur J Biomed Inform.* 2015;11(2)
- [58] Medical Subject Headings, available from <https://nlk.cz/pro-knihovny/mesh/> (cited 20.5.2017)
- [59] SNOMED CT, available from <https://www.nlm.nih.gov/healthit/snomedct/index.html> (cited 19.5.2017)
- [60] iSpell, available from <ftp://ftp.tul.cz/pub/unix/ispell/ispell-czech-20040229.tar.gz> (cited 2009)
- [61] Zvára K, Kašpar V. Identifikace jednotek a dalších termínů v českých lékařských zprávách. *Eur J Biomed Inform.* 2010;6(1)
- [62] Zvára K, Svátek V. Extrahovatelnost informací z českých lékařských zpráv. *Eur J Biomed Inform.* 2012;8(5)
- [63] Přečková P, Zvárová J. The Role of International Nomenclatures and Standards in Travel Shared Health Care. *Travel Health Informatics and Telehealth.* 2009;

- [64] Heid WD, Chasteen J, Forrey AW. The electronic oral health record. *Contemp Dent Pract.* 2002;3
- [65] Zvárová J, Dostálová T, Hanzlíček P, Teuberová Z, Nagy M, Pieš M. Electronic health record for forensic dentistry. *Methods Inf Med.* 2008;47
- [66] Chleborád K, Dostálová T (školitel). Disertační práce: Stav chrupu u hendikepovaných pacientů. Praha: Univerzita Karlova; 2014.
- [67] Chleborád K, Zvára K, Dostálová T, Zvára K, Hippmann R, Ivančáková R et al. Evaluation of voice-based data entry to an electronic health record system for dentistry. *Biocybernetics and biomedical engineering.* 2013;33
- [68] Zvára K, Tomečková M, Peleška J, Svátek V, Zvárová J. Tool-supported Interactive Correction and Semantic Annotation of Narrative Clinical Reports. *Methods Inf Med.* 2017;56(3)
- [69] LOINC, available from <https://loinc.org/> (cited 19.5.2017)
- [70] Databáze léků, available from <http://www.sukl.cz/modules/medication/search.php> (cited 20.5.2017)
- [71] "Málem jsem zabil bratra!" Nový překladač je lepší, chyby jsou vtipnější., available from [http://technet.idnes.cz/google-prekladac-neuronove-site-google-translate-srovnani-test-p80-/sw\\_internet.aspx?c=A170509\\_130949\\_sw\\_internet\\_pka](http://technet.idnes.cz/google-prekladac-neuronove-site-google-translate-srovnani-test-p80-/sw_internet.aspx?c=A170509_130949_sw_internet_pka) (cited 29.5.2017)

- [72] Tange HJ, Hasman A, De Vries Robbe PF, Schouten HC. Medical narratives in electronic medical records. *Int J Med Inf.* 1997;46
- [73] Blobel B, Hasman A, Zvárová J. *Data and Knowledge or Medical Decision Support.* 2013.
- [74] Hoerbst A. *Exploring Complexity in Health: An Interdisciplinary Systems Approach.* 2016.

## *Seznam publikací doktoranda – ve vztahu k tématu disertace*

### *a) s impact factorem*

1. **Zvára K**, Tomečková M, Peleška J, Svátek V, Zvárová J. Tool-supported Interactive Correction and Semantic Annotation of Narrative Clinical Reports. *Methods Inf Med* 2017;56(3) **IF: 1,772**
2. Chleborád K, **Zvára K**, Dostálová T, Zvára K, Hippmann R, Ivančáková R, Zvárová J, Smidl L, Trmal J, Psutka J. Evaluation of voice-based data entry to an electronic health record system for dentistry. *Journal of Biocybernetics and Biomedical Engineering*, 2013, 33(4), 204-210, **IF: 0,234**
3. Přečková P, Zvárová J, **Zvára K**. Measuring diversity in medical reports based on categorized attributes and international classification systems. *BMC Medical Informatics and Decision Making*, 2012, 12: 31, **IF: 1,603**

### *b) bez impact factoru*

1. Chleborád K, **Zvára K**, Dostálová T, Zvára K, Ivančáková R, Zvárová J, Smidl L, Trmal J, Psutka J. Dental Electronic Health Record Evaluation. *EFMI STC Prague Proceedings*. 2013
2. Chleborád K, **Zvára K**, Dostálová T, Ivančáková R, Zvárová J, Smidl L, Trmal J, Psutka J. Dental Electronic Health Record Evaluation. *International Journal on Biomedicine and Healthcare*. 2013, 1(1) s. 50-50
3. Chleborád K, Dostálová T, **Zvára K**, Seydlová M, Zvárová J, Hippmann R. Semantic Interoperability in Dentistry. *Eur J Biomed Inform*, 2012, 8(5), s. 4-5
4. Chleborád K, **Zvára K**, Dostálová T, Seydlová M, Ivančáková R, Zvára K, Zvárová J. Porovnávání záznamu stavu chrupu třemi metodami – elektronická zdravotní dokumentace versus ruční zápis do WHO karty. *Česká stomatologie a Praktické zubní lékařství*. 2011, 59(3), s. 57-64

5. **Zvára K.** Strukturální a lexikální analýza lékařských zpráv. Doktorandské dny '11, Ústav informatiky AV ČR, v.v.i., Matfyzpress 2011, s. 145-148
6. **Zvára K,** Kašpar V. Identification of Units and Other Terms in Czech Medical Records. Eur J Biomed Inform. 2010 6(1)
7. Dostálová T, Zvárová J, **Zvára K,** Seydlová M, Chleborád K, Kašpar V. A Voice Controlled Software Tool Based on Structure Electronic Health Record. Methodological Issues in Oral Health Research: Intervention Studies. Marmara University 2010, s. 37-38
8. Seydlová M, Dostálová T, **Zvára K,** Zvárová J, Feberová J. New Strategy for Teaching Medical Students Stomatology. MEFANET 2009, Brno, s. 1-3
9. Zvárová J, Chleborád K, **Zvára K.** Medical Informatics and Information Technology Supporting Oral Medicine. Health Informatics Meets Ehealth Volume 223. 2016, pp. 230-236
10. Zvárová J, Chleborád K, **Zvára K,** Dostálová T, Studies in health technology and informatics, 2016, vol. 223, pp. 230-236, 2016
11. Zvárová J, Lhotská L, Seidl L, **Zvára K,** Studies in health technology and informatics, 2012, vol. 180, pp. 672-676, 2012
12. **Zvára K,** Svátek V. Extrahovatelnost informací z českých lékařských zpráv. Eur J Biomed Inform. 2012 8(5), s. 43-47

## ***Seznam publikací doktoranda – bez vztahu k tématu disertace***

### ***a) s impact factorem***

1. Zvárová J, Peleška J, Hanzlíček P, **Zvára K**. Enhanced Care of Hypertensive Patients using Internet. Medical Informatics and the Internet in Medicine. 2002, 27(3), pp. 161-168, **IF: 0,698**
2. Martínková P, **Zvára K**, Zvárová J, Zvára K. The new features of the ExaMe evaluation systém and reliability of its fixed tests. Methods Inf Med, 2006, 45(3), pp. 310-315, **IF: 0,97**

### ***b) bez impact factoru***

1. Zvárová J, Hanzlíček P, Hejl H, Jirkovec Z, Pikhart H, Příbík V, Smitková V, **Zvára K**. Základy informatiky pro biomedicínské obory
2. Zvárová J, Peleška J, **Zvára K**, Mareš R. Electronic Health Record and Medical Guidelines in Cardiology. Analysis of Biomedical Signals and Images. 2002, pp. 448-450
3. Peleška J, **Zvára K**, Veselý A, Zvárová J, Tomečková M, Hanzlíček P. Development of Electronic Form of the 1999 WHO/ISH Hypertension Guidelines. Health Data in the Information Society. 2002, pp. 268-271
4. Peleška J, **Zvára K**, Anger Z, Tomečková M, Veselý A, Zvárová J. Elektronická forma lékařských doporučení – usnadní se jejich naučení a používání? Metabolismus. 2002 5(3) pp. 26-27
5. Peleška J, **Zvára K**, Tomečková M, Veselý A, Zvárová J. Elektronická forma lékařských doporučení – příspěvek k jejich lepšímu využívání. Inteligentní systémy ve zdravotní péči. Abstrakty. Praha, ČVUT, 2002, s. 5
6. Zvárová J, Příbík V, Peleška J, Tomečková M, **Zvára K**. Elektronický zdravotní záznam a telemedicína. Inteligentní systémy ve zdravotní péči. Abstrakty. Praha, ČVUT, 2002, s. 8

7. Peleška J, **Zvára K**, Tomečková M, Zvárová J. Computer Assisted Implementation of the 1999 WHO/ISH Hypertension Guidelines. *Journal of Hypertension*, 2002, 20(4), s. 86
8. Peleška J, **Zvára K**, Tomečková M, Veselý A, Zvárová J. Elektronická lékařská doporučení pro léčbu hypertenze. *Cor et Vasa*, 2002, 44(4), s. 90-91
9. Zvárová J, Peleška J, Hanzlíček P, **Zvára K**. Enhanced Care of Hypertensive Patients using Internet. *Technology and Health Care*. 2001, 9(6), s. 487-488
10. Hanzlíček P, **Zvára K**. Internet po lžičkách. *Lékař a technika. Biomedicínské inženýrství a informatika*. 1998, 29(1) s. 21-22
11. Hanzlíček P, **Zvára K**. Internet po lžičkách. *Lékař a technika. Biomedicínské inženýrství a informatika*. 1998, 29(2) s. 49-50
12. Zvárová J, Hanzlíček P, Peleška J, Štefek M, Švejda D, Tomečková M, **Zvára K**, Choura M, Gula P. Education and Training in Medical Informatics, Statistics and Epidemiology at the EuroMISE Center. *Medicinska informatika 97*. Zagreb. 1997. s. 89-91
13. Zvárová J, Hanzlíček P, Peleška J, Štefek M, Švejda D, Tomečková M, **Zvára K**, Choura M, Gula P. Education and Training in Medical Informatics, Statistics and Epidemilogy at the EuroMISE Center. *Telecommunications for Education and Training*. Prague, 1999, s. 176-180
14. Zvárová J, **Zvára K**. Evaluation of Knowledge Using ExaMe Program on the Internet. *User Acceptance of Health Telematics Applications: Education and Training in Health Telemetics*. Amsterdam, 2000. s. 145-151
15. **Zvára K**, Hanzlíček P, Zvárová J. ExaMe – A Program for Evaluation of Medical Knowledge Using the Internet. *MEDNET 98*. Birmingham. 1998, s. 39
16. Hanzlíček P, **Zvára K**. Internet po lžičkách. *Lékař a technika. Biomedicínské inženýrství a informatika*. 1997 28(6), s. 143-144

17. Zvárová J, Hanzlíček P, Gula P, **Zvára K**. Internet and its Role in Medical Education. MEDNET 97. Brighton, 1997, s. 43
18. Martinková P, **Zvára K**, Dostálová T, Zvárová J. Education for Medical Decision Support at EuroMISE Centre. EFMI STC Prague Proceedings, 2013
19. Martinková P, **Zvára K**, Dostálová T, Zvárová J. Education for Medical Decision Support at EuroMISE Centre. International Journal on Biomedicine and Healthcare. 2013 1(1)
20. Zvárová J, Svačina Š, Dostálová T, Seydlová M, **Zvára K**. Multiple Disciplines Synergy Tools for Ph.D. Students of Biomedical Informatics at Charles University in Prague. Quality of Life through Quality of Information. Amsterdam. 2012, s. 1153-1155
21. Zvárová J, Svačina Š, Dostálová T, Seydlová M, **Zvára K**, Papíková V, Zvolský M, Štuka Č, Vejražka M, Feberová J. Biomedical Informatics Education at Charles University in Prague for Undergraduate and Doctoral Degree Studies. Eur J Biomed Inform, 2011, 7(2), s. 72-78
22. **Zvára K**. Vliv ekonomických faktorů na efektivitu zdravotní péče. Konference Health Technology Assessment, Praha, 2010
23. Zvárová J, Dostálová T, **Zvára K**, Heroutová H. Biomedical Informatics Doctoral Programme and Lifelong Education. MEDINFO 2010, s. 1426-1426
24. Zvárová J, Dostálová T, **Zvára K**, Seydlová M, Papíková V, Zvolský M. E-learning and the Impact of Information Delivery to Students and Patients in Multilingual Setting. MEFANET 2010, s. 1-10
25. Zvárová J, Lhotská L, Přibík V, Adášková J, Brechlerová D, Hanzlíček P, Huptych M, Kopecký M, Papíková V, Potůček J, Přečková P, Říha A, Svátek V, Šárek M, Zitová B, **Zvára K**. Data a znalosti v biomedicině a zdravotnictví. Karolinum 2010 (427 s.)
26. Dostálová T, Hlíňáková P, Michalík P, Chleborád K, Seydlová M, **Zvára K**, Zvárová J. Computer-Supported Analysis and Therapy of



- Patients with the TMJ Parafunction. EPA 2009 Innsbruck, s. 52-59
27. Zvárová J, **Zvára K**, Rexová P. Exame – the System for Evaluation of Knowledge Using Internet. Proceedings of the 11th World Congress on Medical Informatics. Edmonton 2004, s. 1935
  28. **Zvára K**, Zvárová J. On New Features of the Exame System. International Joint Meeting EuroMISE 2004, Prague, s. 107
  29. Zvárová J, Svačina Š, Valenta Z, Berka P, Buchtela D, Jiroušek R, Malý M, Papíková V, Peleška J, Rauch J, Vajda I, Veselý A, **Zvára K**, Zvolský M. Systémy pro podporu lékařského rozhodování. Karolinum 2009, 504s
  30. Kalina J, Seidl L, **Zvára K**. System for Selecting Relevant Information for Decision Support. Data Knowledge for Medical Decision Support Volume 187, 2013, pp. 83-87
  31. Zvárová J, **Zvára K**, Moutzoglou A. E-Health Systems Quality and Reliability: Models and Standards. 2011. pp. 18-27
  32. Zvárová J, Hanzlíček P, Nagy M, Přečková P, **Zvára K**, Seidl L, Bureš V, Šubrt D, Dostálová T, Seydlová M. Biomedical Informatics Research for Individualized Life-long Shared Healthcare. 2009. Biocybernetics and Biomedical Engineering, 29(2), pp. 31-41
  33. Nagy M, Hanzlíček P, Deoszegi M, Zvárová J, Přečková P, Seidl L, **Zvára K**, Bureš V, Šubrt D. Applied Information Technologies for Development of Continuous Shared Health Care. Cesnet conference 2008, pp. 131-138